

DOCUMENT RESUME

ED 370 343

EC 303 078

AUTHOR Vanderheiden, Gregg C.
TITLE A Standard Approach for Full Visual Annotation of Auditorially Presented Information for Users, Including Those Who Are Deaf: ShowSounds.
INSTITUTION Wisconsin Univ., Madison. Trace Center.
SPONS AGENCY Apple Computer, Inc., Cupertino, CA.; Information Technology Association of America, Arlington, VA.; National Inst. on Disability and Rehabilitation Research (ED/OSERS), Washington, DC.
PUB DATE [93]
CONTRACT H133E80021
NOTE 8p.; IBM Corporation also provided support.
PUB TYPE Reports - Descriptive (141)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Accessibility (for Disabled); Auditory Stimuli; Computer Networks; *Computer System Design; *Deafness; Design Requirements; *Hearing Impairments; Information Systems; *Input Output Devices; *Visual Aids
IDENTIFIERS *ShowSounds

ABSTRACT

This paper describes a standard cross-platform strategy for providing deaf individuals with access to sounds generated by electronic information systems. The strategy is based on a "ShowSounds" switch built into the basic system architecture. Turning the switch "on" would signal that all information presented auditorially is also to be presented visually. This would apply to both information presented by the operating system and by individual application programs. Information is provided on the background of the ShowSounds innovation, the types of auditory information, use of redundant auditory information as an access strategy, the ShowSounds approach to visual annotation of important auditory information, types of video annotation, applications of ShowSounds, the ShowSounds implementation schedule, and progress to date. The importance of making both computer and information systems vendors aware of the need for such a capability and to build it into the next generation systems is stressed. (DB)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED 370 343

A Standard Approach for Full Visual Annotation of Auditorially Presented Information for Users, Including Those Who Are Deaf: ShowSounds

Gregg C. Vanderheiden
Director, Trace R&D Center
and Industrial Engineering Department
University of Wisconsin-Madison
S-151 Waisman Center
1500 Highland Avenue
Madison, WI 53705-2280
608/262-6966

Abstract

A standard cross-platform strategy for providing access to sounds generated electronic information systems is proposed. This strategy is built around a "ShowSounds switch" or flag which would be built into the basic system architecture. Turning this (software) "switch" on would be a signal from the user that they would like all information being presented auditorially to also be presented visually. This would apply to both information presented by the operating system and by individual application programs. This visual annotation would go beyond just captioning, and would include other visual display as necessary to convey important information being presented auditorially.

Like the current closed captioning on television, the ShowSounds on computers (and information systems) capability would allow users to indicate whether they wanted to have sound annotation visible or not. Individuals who can hear and who are not interested in visual annotation could just leave the ShowSounds switch turned off. However, if an individual had a severe hearing impairment, was in an environment where noise interfered with hearing, or had the sound turned off (as in a library), they could turn on the ShowSounds switch. Once the ShowSounds switch was turned on, the audiovisual material could automatically provide visual annotation of all important auditory information, allowing the program or device to be used effectively even if it could not be heard.

Since computer-based information and multimedia systems have considerably more intelligence and graphic capabilities than television, the types of annotation that could be provided could also be more diverse. For example, with ShowSounds switch in the on position, a program presenting information via speech could automatically begin showing captions as well. Another program using a rising and falling tone to indicate information could simultaneously display a small scale with a rising and falling pointer to correspond to the tone. Beeps or other tones meant to draw one's attention to the screen could be accompanied by a screen blink or other visual event substantial enough to catch one's attention, even if one were looking at the keyboard or elsewhere. As a result, video annotations could be tailored to best present the auditory information in a visual format which was consistent with the program presentation.

Background

Speech and sound are being used more and more in computer and information systems which had previously been strictly visual. The rapid advance of multimedia platforms as well as

EC 303078

audiovisual information kiosks and information systems over the next 5-10 years will see a dramatic change in the use of sound.

In some cases, the use of sound is ornamental in nature, and not essential to the operation of the device. The playing of music in the background may be pleasant, but not required for operation of an information program. In other cases, the information may be redundant with information being presented on the screen. In these cases, the provision of annotation may not be important. However, where the information being presented auditorially is not also available visually, users who cannot hear would be left trying to operate the system without all of the information presented to hearing users. As the use of sound is incorporated into more and more information systems, education systems, and systems for daily living, the problems faced by those individuals who cannot hear will grow. Since these applications will appear in education and employment as well as daily living, access to these systems is essential.

Three Types of Auditory Information

For the purposes of this discussion, let us break the types of information presented auditorially into three major categories:

- 1) *Decorative*: The use of sounds to enhance the "appearance" or presentation, but which is not required for the operation of the system or device.
- 2) *Redundant*: Information which is presented auditorially, but which is redundant with the information presented visually, and therefore not important for the operation of the system.
- 3) *Important*: information which is only presented auditorially, and which is important to the operation of the system.

Types 1 and 2 present no particular problem for individuals who are having difficulty hearing (including those with hearing impairments and those in noisy environments). It is access to Type 3 information that this paper is directed toward.

Use of Type 2 Auditory Information as an Access Strategy

Before discussing the proposed solution strategy for providing access to Type 3 auditory information, it is important to note the merits in maximizing the Type 2 auditory information; that is, maximizing the amount of auditory information which is also provided visually to all users. Not only does the simultaneous presentation of information in both auditory and visual format eliminate the auditory access problem, if properly done it can also result in more effective presentation of the information to all users than auditory presentation alone. Even if a ShowSounds switch is available and supported by a piece of application software, the more that the application uses simultaneous audio and visual presentation of information for all users, the less special programming or provisions will be needed to accommodate users who must use the ShowSounds feature.

Access to Type 3 Auditory Information

Past Approaches

Where information cannot be effectively presented both auditorially and visually, some mechanism needs to be provided to allow individuals who cannot perceive the auditory information to still be able to use the software / system. In the past, auditory information has

largely been limited, particularly on the IBM compatible computers, to beeps or other simple alerting sounds. To allow individuals who are deaf to be aware of these warning sounds, products such as SeeBeep and the ShowSounds feature in AccessDOS have been developed. These products involve the use of terminate-and-stay-resident (TSR) code which directly monitors the data lines that drive the speaker. Whenever activity on these lines is detected (which happens whenever a sound is being made at the speaker), the screen is blinked, or some other visual event is created. As long as these events are simple, undifferentiated warning sounds, this strategy works well. [This feature is hereinafter referred to as "SoundSentry" – a subfunction of the ShowSounds concept.]

This "hardwired" ShowSentry approach to connecting a visual event to any action of the speaker has the advantage that it does not require any cooperation on the part of the application software developers. It will detect any activity on the speaker and cause a screen flash. As long as software never generates any differentiated tones or other sounds, this approach can be quite effective. If the application program uses different tones to provide different information, however, the technique fails. Also, if the program generates decorative sounds (music, or other nonessential sounds), the system will provide screen flashing or other visual events which are both unnecessary, uninterpretable and perhaps annoying. If the program is generating speech, then the system similarly will generate flickering visual events which convey no useful information. Finally, if the program uses some other sound channel besides the built-in speaker to present the auditory information, it would be totally missed by this hardwired approach.

For simple DOS programs which have in the past generally confined themselves to undifferentiated warning beeps, this "hardwired SoundSentry" approach is useful. However, for the next generation of computer programs as well as multimedia computing and information systems, this brute force approach will be of little use. (NOTE: In the AccessDOS program, only the SoundSentry subfunction of ShowSounds has been implemented to date. This subfunction of ShowSounds should not be confused with the general, full ShowSounds concept being presented in this paper.)

A Cooperative Approach to Visual Annotation of Auditory Information

For the next generation of multimedia software and information systems, it will not be possible for the operating system by itself to effectively generate all of the visual cues necessary to deal with the rich variety of auditory information to be presented by applications. As discussed above, much of the auditory information (e.g., *decorative* or *redundant* auditory information) may in fact not require any additional visual treatment at all. However, that auditory information which is both important and presented only in the auditory channel would need to be made accessible. This process would include a) identifying and singling out this information from the nonessential or redundant auditory information, b) determining an appropriate and effective visual equivalent, and c) determining the proper way to present the visual information without interfering with the other visual information on screen. The only people who can do this effectively are the programmers who are actually writing the application software in the first place. When programming the auditory portions, they could either accompany important audio information with visual presentation for all users, or they could provide special cues (captions or other techniques) along with the regular auditory information which users would see on request. The programs would then need some "switch" to know whether the regular or the visually enhanced form should be presented.

System-Wide versus Application Specific Switch

This "Please Show Visual Annotations" switch could be built into each piece of software individually. The problem here is that an individual who required visual cues would have to

keep finding the switch in each particular application program and turning it on. A much better approach would be to have a system-wide "ShowSounds" flag/switch. Such a switch would normally be included in the control panel for the system, in the same place that volume, keyboard characteristics, etc., were adjusted. By flipping the ShowSounds switch on, a flag would be set up which all application programs could easily check while they were running. In this fashion, a single switch could be used to provide an indication which all programs could use.

Multiple Access Switch

In some computers, there may be multiple control panels. For example, there may be a standard computer control panel, and then a separate multimedia control panel. Similarly, it is possible that there be a control panel dealing with sound, and a separate control panel which provides a variety of disability access features. In this case, the question may arise as to the best location for the ShowSounds switch. The recommended location would be wherever other sound information, such as volume, is controlled. If volume is controlled from more than one location, then the ShowSounds switch could also be located in more than one control panel. Like a stairway light which can be controlled from more than one switch, it is very simple to have a system switch/flag which can be operated from different control panels.

Types of Video Annotation

When thinking of video presentation of auditory information, one of the first things to come to mind is the use of captions. Captions can be an especially effective tool. They can be used not only to present speech, but also to describe sounds. Captions alone, however, are not sufficient. For beeps or warning sounds, some type of screen blink other visual event major enough to catch one's attention, even if one is looking at the keyboard, may be necessary. This may involve a blink of the full screen, or just the bottom edge of the screen, which is generally still within the visual field when an individual is looking at the keyboard or desktop. For other types of auditory information, however, it may be necessary to have a real-time qualitative feedback. For example, information which is conveyed by a continually changing pitch cannot be presented effectively either as a caption or a flash. In this case, some type of pointer on a scale which moved up and down in real time to indicate the pitch may be necessary. Again, since the visual presentation is in the hands of the programmers creating the program, the visual event can be tailored to meet the particular needs of the program. Conversely, the program can also be designed in such a way that the visual events are easy to incorporate. In fact, it will probably be the case that the visual presentation which began as being a "ShowSounds" means of presenting the auditory information may often be directly incorporated into the normal presentation of the information.

System Support for Visual Annotation

The only thing really required on a system level to support this ShowSounds concept will be the creation of a flag which could be read from any application program, and a switch in the control panel that could be used to set or unset the flag. However, it is very common for operating system to facilitate the development and standard presentation of information by providing tools. Similarly, it would be useful for operating systems to include certain tools that would facilitate captioning and other visual annotation techniques. With enough study, there may in fact evolve a series of human interface guidelines specifically covering auxiliary visual presentation of auditory information. These guidelines, combined with the system tools, could greatly facilitate the work of application and information system developers in the creation of effective and consistent program interfaces in this area.

Although work has only begun in this area, an initial breakdown of the types of auditory information might be:

- 1) Speech,
- 2) Other sound events which can be easily and quickly described in words,
- 3) Undifferentiated warning beeps or tones,
- 4) Differentiated warning beeps or tones,
- 5) Continuous scalar auditory information (e.g., a varying pitch), and
- 6) Information not directly visually presentable (e.g., music).

Some system tools might be:

- Automated visual presentation of system beep functions,
- Sizable, locatable caption boxes or bubbles,
- Autocaptioning (for systems with built-in speech synthesizers), and
- Undifferentiated and differentiated alert utilities.

Any tools that involve the autocaptioning of synthesized speech should also have the ability to turn the captioning off via data in the text stream, in order to avoid captioning of the speech when it is redundant with what is already on the screen.

Application of ShowSounds

The ShowSounds concept, then, consists of a system-wide switch or flag which a user can set when they would like to have a visual presentation of any important auditory information. Individual application programs can then be written to support ShowSounds or not. If the operating system is written to automatically provide a visual indication of its system beep function, then programs which use the system beep as an alert would not need to worry about their own visual presentation. Otherwise, programs not supporting ShowSounds would not provide any visual cues. Programs which choose to be "ShowSounds compatible" would look for the status of the ShowSounds flag and provide captioning or other visual presentations as and when necessary. Such programs could be marked on their boxes and manuals with a small symbol and/or the phrase "ShowSounds compatible" to indicate this. Schools, government agencies, and companies interested in having software, training programs, etc., which were accessible to their older or younger employees with hearing impairments could look for such an indication on the products. Particularly where there are multiple products on the market providing basically the same function, these types of provisions could effect purchasing decisions. To the extent that such compatibility and accommodations were considered in the purchasing process by schools, governments, etc., these considerations would provide the incentive to developers to incorporate ShowSounds compatibility in their application programs.

SoundSentry

Since not all programs will choose to be ShowSounds compatible, there is still be a role for a system-level "SoundSentry" function as an auxiliary or subfunction of ShowSounds. This would be a function which could be turned on by the user, and which would provide a visual indication of all sounds being created. Such a SoundSentry should provide both a visual indication that a

sound is being made, and some type of an analog indication of the pitch if possible. Typically, this SoundSentry feature would be turned on for particular programs, and then turned off again when working with ShowSounds-compatible programs, or programs which use nothing but system beeps. The SoundSentry would not be anywhere near as effective as a properly designed ShowSounds-compatible program, but it would provide the user with some type of visual indication of the existence and character of the sounds being generated. For sounds other than speech, this might be an effective aid to using numbers of uncooperative programs.

ShowSounds Implementation Schedule

In some respects, ShowSounds is a chicken-and-egg proposition. There is no value to have a ShowSounds flag in the system if there are no applications that use it. On the other hand, there will never be applications that use it if the system flag does not exist. Fortunately, the creation of the ShowSounds flag on a system level is relatively simple. In addition, if it is tied to the system beep, it can have immediate impact on providing visual notification of all system beep events. For many applications at the present time, the only sounds that are made are generated using the system beep.

A schedule of implementation of ShowSounds may therefore look like the following:

- 1) Major operating system vendors incorporate a ShowSounds flag and put a switch in one or more of their system control panels to provide central access by users to the flag.
- 2) Operating systems adapt their system beep function to check for the ShowSounds flag, and provide a visual indication whenever the system beep is triggered by an application.
- 3) The implementation of a SoundSentry function may also be considered. However, this approach may be fairly ineffective with multi-media systems which have several different custom sound generating cards, etc.
- 4) Application developers support the ShowSounds flag by checking its setting and providing visual presentations for any auditory information. Programs that have little or no auditory information, or which only use system beeps, would be easily be able to qualify and label themselves as ShowSounds-compatible.
- 5) The operating systems manufacturers provide closed captioning and other visual annotation tools to facilitate ShowSounds implementation by third-party developers.
- 6) Autocaptioning capabilities in connection with any system-based text-to-speech utilities are provided.

Progress to Date

Some progress has already been made toward the goal of producing support for the ShowSounds capability across company and product lines. Discussions are underway with Apple Computer, Microsoft, and IBM regarding the inclusion of the ShowSounds switch as a part of the standard control panel for sounds in future versions of their operating systems.

The SoundSentry subfunction of ShowSounds has been incorporated into the AccessDOS software package which is distributed by IBM as a complement to its DOS operating system. AccessDOS provides several important disability access features for IBM PC and PS/2 computers running DOS. (In AccessDOS, the SoundSentry/ShowSounds switch is currently tied

to a feature which causes either a small note in the upper left-hand corner of the screen, or a full screen blink, whenever a sound of significant length is initiated by the computer.)

The ShowSounds (SoundSentry) feature with accompanying screen blink is also included in the current release of the Access Utility being distributed by Microsoft for its Windows operating system. In addition, Microsoft has urged developers to check the WIN.INI file for the existence of a ShowSounds flag, and to support full ShowSounds in their applications.

The full ShowSounds concept is included in the White Paper which was commissioned by the Information Technology Foundation (ITF) (formerly the ADAPSO Foundation). The White Paper is the first step toward the development of design guidelines for the applications software industry. ITF is a non-profit foundation of the Information Technology Association of America, the trade association for most of the major application software companies in America. The guidelines are being developed in response to a request by industry for direct, industry-usable information as to how software products can be designed to be more accessible to people with disabilities.

These efforts to develop the ShowSounds concept are also being coordinated with the Caption Center, a service of WGBH Education Foundation. The Caption Center, through its Media Access Research and Development Office, is leading a multifaceted effort to provide access to all types of media: television, radio, print, live performance, theatrical film, and others.

Conclusion

Initial efforts have begun toward the implementation and support for the ShowSounds feature on a number of computer platforms. Much work is needed, however, to make both computer and more general information system vendors aware of the need for such a capability, and to build it into the next generation systems.

Acknowledgements

This work has been supported in part by Grant H133E80021, from the National Institute on Disability and Rehabilitation Research of the US Department of Education, and by Apple Computer Corporation, IBM Corporation, and the Information Technology Foundation (ITF) of the Information Technology Association of America (ITAA), formerly ADAPSO.

For more information, contact Gregg C. Vanderheiden, Ph.D., Trace R&D Center, S-151 Waisman Center, 1500 Highland Avenue, Madison WI 53705, voice 608/262-6966, fax 608/262-8848, e-mail vanderhe@mac.wisc.edu.